REMARKS

Applicants respond hereby to the outstanding Office Action mailed on August 31, 2007. Claims 21-27 are amended hereby, and new claims 28-30 are presented for examination; amended claims 21, 25 and 26, and newly presented claim 28 are the pending independent claims. Moreover, applicants thank Examiner Bruckart for his efforts in working with them to realize claims that effectively communicate and limit the subject matter of their invention as set forth in the application as filed. No new matter is added.

In the August 31, 2007, Office Action, the Examiner rejected claims 21-24, 25 and 26 under 35 U.S.C. 102 as anticipated by U.S. Patent 6,178,529 to Short, et al. (Short), and rejected claim 27 under 35 USC 103(a) as unpatentable over Short in view of U.S. Patent 5,761,506 to Angle, et al. (Angle). Claims 21, 25 and 26 were also rejected under 35 U.S.C. 112 as being indefinite.

For the reasons discussed below, claims 21-27 as amended, and new claims 28-31, are clear and definite and fully comply with the requirements of 35 U.S.C. 112. Each of amended claims 21-27, and new claims 28-31 patentably distinguish over the prior art and are allowable, at least for the reasons set forth below. The Examiner is, thus, respectfully asked to reconsider and to withdraw the rejection of claims 21-26 under 35 U.S.C. 102, the rejection of claim 27 under 35 U.S.C. 103(a) and the rejections of claims 21, 25 and 26 under 35 U.S.C. 112, and to allow each of amended claims 21-27 and new claims 28-31.

Response to Rejections under 35 USC §112, Second Paragraph

In rejecting claims 21, 25 and 26 under 35 U.S.C. 112, the Examiner asserts that the recitation of "a globally optimal configuration" is unclear, and how the configuration is globally optimal. The Examiner also asserts that the use of the phrase "occasionally changing resources" is indefinite, particularly in that the phrase does not make clear how a first group of static and occasional changing resources differs from a group of dynamically changing events. The Examiner further asserts that the use of the phrase "including, for each service, ... " lacks antecedent support, that the phrase "the supporting services" lacks antecedent support, and that the phrase "a desired level of automation" is unclear.

Applicants respond hereby to the rejections under Section 112, and the Examiner's comments in the Examiner's Remarks as stated in the first three lines of page 9 of the outstanding Office Action. That is, applicants have amended each of the pending claims to address not only the Examiner's specific assertions, but to clean up and resolve and possible or inherent Section 112, second paragraph issues with the claims, to render the inventive subject matter definite and clear in those claims. The claims are substantially modified, and the inventive subject matter as recited by the amended claim language renders the claims clear, and more readily distinguishable from the prior art cited in the Section 102(e) and Section 103(a) rejections addressed below.

Claim 21 (as amended), for example, recites building an optimal cluster configuration of networked resources, said optimal cluster configuration optimized in view of a current state of availability and quality fo service and accordance with the rule-based constraints. The step of building then qualifies the availability and quality of service. The claim 21 method manages the cluster based on the availability and quality of service for each of the resources at the time the cluster is built, or constructed. Because during cluster operation, the resources comprising the cluster at any given point may fail, go off-line and come on-line, and because need for or load on any given resource may dynamically change, instantly, or over time, the current cluster configuration state dynamically changes over time, and can migrate away from the (built) global optimal cluster configuration.

In view of the dynamic dependencies of and configuration information about the cluster, the current detected cluster or resources state and occurrence of system events during cluster operation, the current cluster is modified by allocating and reallocating the resources to match or rebuild a new or alternative globally optimal cluster configuration, which is constructed based on the rule-based constraints, and the current cluster resources. Rebuilding or realigning during cluster operation occurs only when needed, where need is qualified by the rule-based constraints. Amended claims 21, 25 and 26 use of the phrase globally optimal cluster configuration is believed to be definite, and otherwise comply with the second paragraph of section 112. For that matter, the reference to service in lieu of the more proper term "resource" is now obviated in view of the claims amendments as well as the phrase "desired level of automation."

Applicants respectfully submit, therefore, that amended claims 21-27 are clear and definite and fully comply with the requirements of 35 U.S.C. 112. Consequently, the Examiner is asked to reconsider and to withdraw the rejections of Claims 21, 25 and 26 under 35 U.S.C. 112.

Response to Rejection Under 35 USC §102

Claims 21-24, 25 and 26 were rejected under 35 USC §102(e) in view of Short. The Examiner asserts that Short discloses a method of managing a cluster of networked resources using rule-based constraints in a scalable clustering environment (col. 5, line 46-col. 6, line 9), the method comprising the steps of:

building a globally optimal configuration of said cluster of resources(col. 5, line 46-col. 6, line 9; col. 7, lines 38-54), wherein each of the resources has an availability and quality of service, and the availabilities and quality of services of the resources are determined by dependencies among the resources, user preferences, constraints on the resources, events, and network policies (col. 6, lines 28-45; col. 8, line 42-50),

bringing said cluster of resources on-line in a systematic manner (col. 5, lines 46-53, col. 6, lines 46-65), given current states of said resources and resource groups, and said dependencies, preferences, constraints, events, and policies (col. 5, lines 23-53);

determining dynamic dependencies of and configuration information about said cluster of resources, including determining said dependencies and configuration information (i) at cluster initialization (col. 5, lines 46-53, col. 6, lines 46-65) and (ii) dynamically during cluster operation (col. 7, lines 13-53),

supporting startup and shutdown of said cluster of resources according to current policies, and system events (col. 5, lines 46-53),

separating said dependencies, preferences, constraints, events, and policies into (i) a first rules based group (col. 5, line 46-col. 6, line 9) and (ii) a second, dynamically changing events based group (col. 6, lines 28-45), wherein said first group captures the static resources and the occasionally changing resources, including, for each service, the type and quality of the supporting services needed to enable said each service(col. 5, line 46-col. 6, line 9; col. 6, lines 46-65), and

combining said first and second groups in a systematic manner only when needed to build the said optimal configuration, whereby by separating the dynamic dependencies, preferences, constraints, events and policies from other dependencies, preferences, constraints, events and policies, and then combining said first and second groups in a systematic manner only when needed, a desired level of automation is achieved in the coordination and mapping of resources and services (col. 5, lines 22-36).

Independent claims 25 and 26 are rejected on substantially the same grounds.

As discussed above with respect to the rejections under Section 112, second paragraph, the invention provides a method (and system) for managing a cluster of networked resources using rule-based constraints in a scalable clustering environment. Independent claim 21 as amended sets forth a method of managing a cluster of networked resources and resource groups using rule-based constraints in a scalable clustering environment, the method comprising the steps of:

building a globally optimal cluster configuration of said networked resources in accordance with said rule-based constraints and a current state of said resources, including identifying for each of the resources and resource groups an availability and quality of service, which are determined by dependencies among the resources and resource groups, user preferences, constraints on the resources and network policies;

bringing said cluster of networked resources on-line in a systematic manner, given the current state of each of the resources and resource groups, and their equivalency, dependencies, user preferences, constraints on the resources, and network policies,

with said cluster of networked resources on-line, determining dynamic dependencies of and configuration information about said cluster of networked resources (i) statically at said step of building and said step of bringing said cluster of networked services online and (ii) dynamically during cluster operation in accordance with said rule-based constraints,

supporting startup, operation and shutdown of said cluster of networked resources according to current policies, and system events, and said rule-based constraints;

separating said dependencies among resources and resource groups, user preferences, constraints among the resources, system events, and current policies into (i) a first, static rules based group and (ii) a second, dynamically changing events based group, wherein said first group captures the static resources, including, for each resource, a type and quality of the supporting resources needed to enable said each resource, wherein said step of separating is implemented according to said rule-based constraints; and

combining said first and second groups in a systematic manner only when needed to build the said globally optimal cluster configuration, or an alternative globally optimal cluster configuration only when needed during operation to modify and realign the current state of said

cluster to said globally optimal cluster configuration in view of said current policies, said system events and said rule-based constraints.

With this invention, resource availability and quality of service is dependent on the availability and the quality of service provided by one or more other resources in the cluster, and equivalency of the resources or groups of resources. In claim 21 as amended, the dependencies of the networked resources are separated in two groups, where the first group captures the static resources, such as the type and quality of service of the supporting resource needed to fully enable the resource. The second group captures the dynamic state of the various resources, where dynamic changes are captured as system events.

The present invention, by separating the dynamic part (the events) from other parts (the rules) according to said rule-based constraints and combining said first and second groups in a systematic manner only when needed to build the said globally optimal cluster configuration, or alternative optimal cluster configuration only when needed during operation to modify and realign the current state of said cluster to said globally optimal cluster configuration in view of said current policies, said system events and said rule-based constraints provides for managing applications and resources in a scalable, policy-critical clustering environment not found in Short. For that matter, Short does not teach or suggest that the resources may be interchanged with resources, or resource group equivalencies, which is an important feature of the invention.

Short discloses a method and system to facilitate the control and monitoring of disparate resources. With the procedure disclosed in Short, a resource component is connected to a

resource object for management of that object, and a resource monitor connects the resource components to a cluster service. The resource component includes a plurality of methods, and these methods are called by the resource monitor to control and monitor operation of the resource object through the resource component.

While the Examiner states that Short's col. 5, line 46 – column 6, line 9, discloses separating the network resources, resource groups, cluster configurations into static and dynamically changing groups, applicants respectfully disagree. This cited portion of Short. discusses the operation of the resource manager 86 and how that manager makes management decisions and initiates appropriate actions, such as startup, restart and failover. Various factors may be taken into account when making these decisions, but there is no disclosure or teaching of <u>separating</u> the resources and resource groups, or the factors that determine the availability and quality of services of the resources, into the two above-identified groups that are used in the practice of the present invention. Instead, with the procedure of Short, the resource manager receives resource and system state information from a resource monitor and a node manager, and uses that information to make decisions.

And while the Examiner asserts that Short at col. 5, lines 22-36, describes combining the first and second groups only when needed to build the globally optimal configuration, applicants have amended the independent claims in this respect to read combining said first and second groups in a systematic manner only when needed to build the said globally optimal cluster configuration, and only when needed during operation to modify and realign the current state of said cluster to said globally optimal cluster configuration in view of said current policies, said

system events and said rule-based constraints provides for managing applications and resources in a scalable, policy-critical clustering environment, applicants do not find same in Short.

Short at col. 5, lines 22-36, describes database manager 80 as implementing functions to maintain a cluster configuration database on a local device and a configuration database 82 on common persistent storage devices, where the database maintains information about physical and logical entities in the cluster. Nowhere does the cited text mention combining said first and second groups in a systematic manner only when needed to build the said globally optimal cluster configuration, and only when needed during operation to modify and realign the current state of said cluster to said globally optimal cluster configuration in view of said current policies, said system events and said rule-based constraints.

In light of the above-discussed differences between independent claims 21, 25 and 26 and Short, and because of the advantages associated with these differences, these claims are not anticipated by Short under section 102(e). Claims 22-24, which depend from claim 21, are patentable therewith. Accordingly, claims 21- 26 patentably distinguish over Short, and applicants respectfully request withdrawal of the rejection of claims 21-26 in view of Short under section 102(e), and allowance of the claims.

<u>Response To Rejection Under 35 USC §102</u>

Claim 27 was rejected under 35 USC §103(a) over Short in view of Angle. The Examiner asserts that Short teaches the method according to claim 24, which depends from the method of independent claim 21, but that Short fails to teach the claim 27 limitations including

the use of entry queues, but that Angle does and it would have been obvious to combine Angle with Short to realize the invention of claim 27. Applicants respectfully disagree.

Angle. describes a procedure for handling cache misses in a computer system. In this procedure, a computer system architecture schedules processes for execution on one of multiple processors, migrates processes between the processors. Also, processes are rescheduled upon the occurrence of a cache miss. There is no teaching in this reference of equivalency, or of separating dependencies, preferences, constraints, events, and policies into two groups and then combining these two groups as is done in the present invention. The references to queues found at col. 10, lines 41-46, describes prior art cache misses, and how Angle, when experiencing a cache miss in the pre-fetch unit 113, the memory processor 180 execute a read of the missing cache line in memory while the process is in input queue 181. The queue described by Angle is not equivalent to the postprocessor queue as claimed.

In addition to the differences between Angle's queue and applicants' queue as claimed, applicants do not agree that Short includes each of the elements comprising dependent claim 24, and independent claim 21 from which it depends, for at least the reasons set forth above in response to the section 102(e) rejections in view of Short. In light of the above-discussed differences between claims 21 and 24 in view of Short, and the differences between the cited Angle text and claim 27, and because of the advantages associated with these differences, claim 27 is not obvious in view of Short and Angle, and applicants respectfully request withdrawal of the rejection.

For the reasons explained above, the Examiner is respectfully requested to reconsider and to withdraw the rejection of Claims 21-26 under 35 U.S.C. 102(e), and the rejection of claim 27 under 35 U.S.C. 103(a). The Examiner is also asked to reconsider and to withdraw the rejection of Claim 21, 25 and 26 under 35 U.S.C. 112, and to allow Claims 21, 25 and 26. If the Examiner believes that a telephone conference with applicants' undersigned attorney would be advantageous to the disposition of this case, the Examiner is asked to telephone the undersigned.

Respectfully Submitted,

John F. Vodopia
Registration No. 36, 299
Attorney for Applicants

Scully, Scott, Murphy & Presser, P.C.
400 Garden City Plaza – Suite 300
Garden City, New York 11530
(516) 742-4343

JFV:tb